

Rec'd PCT/PTO 01 MAR 2006

10/537612

**RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/537,612  
Source: PCP  
Date Processed by STIC: 3/1/06

***ENTERED***



PCT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/537,612

DATE: 03/01/2006

TIME: 13:55:57

Input Set : A:\Q88296 - 10-537,612 sequence listing.txt  
 Output Set: N:\CRF4\03012006\J537612.raw

```

3 <110> APPLICANT: Imperial College Innovations Ltd
5 <120> TITLE OF INVENTION: Engineering Redox Proteins
7 <130> FILE REFERENCE: Q88296
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/537,612
C--> 10 <141> CURRENT FILING DATE: 2005-06-03
12 <160> NUMBER OF SEQ ID NOS: 11
14 <170> SOFTWARE: PatentIn Ver. 2.1
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 84
18 <212> TYPE: DNA
19 <213> ORGANISM: Escherichia coli
21 <220> FEATURE:
22 <221> NAME/KEY: CDS
23 <222> LOCATION: (1)..(84)
24 <223> OTHER INFORMATION: Helix 1 of E.coli repressor of primer (rop)
26 <400> SEQUENCE: 1
27 acc aaa caa gaa aaa acc gcc ctt aac atg gcc cgc ttt atc aga agc      48
28 Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe Ile Arg Ser
29      1           5           10          15
30 cag aca tta acg ctt ctg gag aaa ctc aac gag ctg                         84
31 Gln Thr Leu Thr Leu Leu Glu Lys Leu Asn Glu Leu
32      20          25
33 <210> SEQ ID NO: 2
34 <211> LENGTH: 28
35 <212> TYPE: PRT
36 <213> ORGANISM: Escherichia coli
37 <400> SEQUENCE: 2
38 Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe Ile Arg Ser
39      1           5           10          15
40 Gln Thr Leu Thr Leu Leu Glu Lys Leu Asn Glu Leu
41      20          25
42 <210> SEQ ID NO: 3
43 <211> LENGTH: 84
44 <212> TYPE: DNA
45 <213> ORGANISM: Escherichia coli
46 <400> SEQUENCE: 3
47 gat gaa cag gca gac atc tgt gaa tcg ctt cac gac cac gct gat gag      48
48 Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu His Asp His Ala Asp Glu
49      1           5           10          15

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/537,612

DATE: 03/01/2006

TIME: 13:55:57

Input Set : A:\Q88296 - 10-537,612 sequence listing.txt  
 Output Set: N:\CRF4\03012006\J537612.raw

```

65 ctt tac cgc agc tgc ctt gcc cgt ttc ggc gac gac          84
66 Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Asp Asp
67      20           25
70 <210> SEQ ID NO: 4
71 <211> LENGTH: 28
72 <212> TYPE: PRT
73 <213> ORGANISM: Escherichia coli
75 <400> SEQUENCE: 4
76 Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu His Asp His Ala Asp Glu
77   1           5           10           15
79 Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Asp Asp
80      20           25
84 <210> SEQ ID NO: 5
85 <211> LENGTH: 192
86 <212> TYPE: DNA
87 <213> ORGANISM: Escherichia coli
89 <220> FEATURE:
90 <221> NAME/KEY: CDS
91 <222> LOCATION: (1)..(192)
92 <223> OTHER INFORMATION: wild type dimeric rop
94 <400> SEQUENCE: 5
95 atg ggt acc aaa caa gaa aaa acc gcc ctt aac atg gcc cgcc ttt atc  48
96 Met Gly Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe Ile
97   1           5           10           15
99 aga agc cag aca tta acg ctt ctg gag aaa ctc aac gag ctg gac gcg  96
100 Arg Ser Gln Thr Leu Thr Leu Glu Lys Leu Asn Glu Leu Asp Ala
101      20           25           30
103 gat gaa cag gca gac atc tgt gaa tcg ctt cac gac cac gct gat gag  144
104 Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu His Asp His Ala Asp Glu
105      35           40           45
107 ctt tac cgc agc tgc ctt gcc cgt ttc ggc gac gac ggt gaa aac ctg  192
108 Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Asp Asp Gly Glu Asn Leu
109      50           55           60
112 <210> SEQ ID NO: 6
113 <211> LENGTH: 64
114 <212> TYPE: PRT
115 <213> ORGANISM: Escherichia coli
117 <400> SEQUENCE: 6
118 Met Gly Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe Ile
119   1           5           10           15
121 Arg Ser Gln Thr Leu Thr Leu Glu Lys Leu Asn Glu Leu Asp Ala
122      20           25           30
124 Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu His Asp His Ala Asp Glu
125      35           40           45
127 Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Asp Asp Gly Glu Asn Leu
128      50           55           60
132 <210> SEQ ID NO: 7
133 <211> LENGTH: 384
134 <212> TYPE: DNA

```

RAW SEQUENCE LISTING DATE: 03/01/2006  
 PATENT APPLICATION: US/10/537,612 TIME: 13:55:57

Input Set : A:\Q88296 - 10-537,612 sequence listing.txt  
 Output Set: N:\CRF4\03012006\J537612.raw

```

135 <213> ORGANISM: Artificial Sequence
137 <220> FEATURE:
138 <223> OTHER INFORMATION: Description of Artificial Sequence: Monomeric rop
139 containing all 4 helices in one continuous
140 sequence
142 <220> FEATURE:
143 <221> NAME/KEY: CDS
144 <222> LOCATION: (1)..(384)
145 <223> OTHER INFORMATION: Monomeric rop consisting of helices 1-1'-2'-2 and
146 helices 1 and 1', and 2' and 2 are connected by
147 GGGG loops
149 <400> SEQUENCE: 7
150 atg ggt acc aaa caa gaa aaa acc gcc ctt aac atg gcc cgc ttt atc 48
151 Met Gly Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe Ile
152 1 5 10 15
154 aga agc cag aca tta acg ctt ctg gag aaa ctc aac gag ctg ggt ggc 96
155 Arg Ser Gln Thr Leu Thr Leu Glu Lys Leu Asn Glu Leu Gly Gly
156 20 25 30
158 ggt ggc acc aaa caa gag aag acc gcc ctt aac atg gcc cgc ttt 144
159 Gly Gly Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe
160 35 40 45
162 atc aga tct cag aca tta acg ctt cta gag aag ctt aac gag ctc ggg 192
163 Ile Arg Ser Gln Thr Leu Thr Leu Glu Lys Leu Asn Glu Leu Gly
164 50 55 60
166 gcg gat gaa cag gca gac ata tgt gaa tcg ctt cac gac cac gct gat 240
167 Ala Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu His Asp His Ala Asp
168 65 70 75 80
170 gag ctt tac cgc agc tgc ctt gcc cgt ttc ggt ggc ggt ggc gcg 288
171 Glu Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Gly Gly Gly Ala
172 85 90 95
174 gat gaa cag gca gac atc tgt gaa tcg ctt cac gac cac gct gat gag 336
175 Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu His Asp His Ala Asp Glu
176 100 105 110
178 ctt tac cgc agc tgc ctt gcc cgt ttc ggc gac gac ggt gaa aac ctg 384
179 Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Asp Asp Gly Glu Asn Leu
180 115 120 125
183 <210> SEQ_ID NO: 8
184 <211> LENGTH: 128
185 <212> TYPE: PRT
186 <213> ORGANISM: Artificial Sequence
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Description of Artificial Sequence: Monomeric rop
190 containing all 4 helices in one continuous
191 sequence
193 <400> SEQUENCE: 8
194 Met Gly Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe Ile
195 1 5 10 15
197 Arg Ser Gln Thr Leu Thr Leu Glu Lys Leu Asn Glu Leu Gly Gly
198 20 25 30
  
```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/537,612

DATE: 03/01/2006

TIME: 13:55:57

Input Set : A:\Q88296 - 10-537,612 sequence listing.txt  
 Output Set: N:\CRF4\03012006\J537612.raw

200 Gly Gly Gly Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe  
 201 35 40 45  
 203 Ile Arg Ser Gln Thr Leu Thr Leu Leu Glu Lys Leu Asn Glu Leu Gly  
 204 50 55 60  
 206 Ala Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu His Asp His Ala Asp  
 207 65 70 75 80  
 209 Glu Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Gly Gly Gly Ala  
 210 85 90 95  
 212 Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu His Asp His Ala Asp Glu  
 213 100 105 110  
 215 Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Asp Asp Gly Glu Asn Leu  
 216 115 120 125  
 220 <210> SEQ ID NO: 9  
 221 <211> LENGTH: 19  
 222 <212> TYPE: DNA  
 223 <213> ORGANISM: Artificial Sequence  
 225 <220> FEATURE:  
 226 <223> OTHER INFORMATION: Description of Artificial Sequence: psp7 upstream  
 227 amplification sequence  
 229 <400> SEQUENCE: 9  
 230 gcgaaaattaa tacgactca 19  
 233 <210> SEQ ID NO: 10  
 234 <211> LENGTH: 23  
 235 <212> TYPE: DNA  
 236 <213> ORGANISM: Artificial Sequence  
 238 <220> FEATURE:  
 239 <223> OTHER INFORMATION: Description of Artificial Sequence: asp4  
 240 downstream amplification sequence  
 242 <400> SEQUENCE: 10  
 243 gtggctgt gccaccgctg agc 23  
 246 <210> SEQ ID NO: 11  
 247 <211> LENGTH: 128  
 248 <212> TYPE: PRT  
 249 <213> ORGANISM: Artificial Sequence  
 251 <220> FEATURE:  
 252 <223> OTHER INFORMATION: Description of Artificial Sequence: RDM14.5  
 254 <400> SEQUENCE: 11  
 255 Met Gly Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe Ile  
 256 1 5 10 15  
 258 Arg Ser Gln Thr Leu Thr Leu Leu Glu Lys Leu Asn Glu Leu Gly Gly  
 259 20 25 30  
 261 Gly Gly Gly Thr Lys Gln Glu Lys Thr Ala Leu Asn Met Ala Arg Phe  
 262 35 40 45  
 264 Ile Arg Ser Gln Thr Leu Thr His Leu Glu Lys Leu Asn Glu Leu Gly  
 265 50 55 60  
 267 Ala Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu Ala Asp Trp Ala Asp  
 268 65 70 75 80  
 270 Glu Leu Tyr Arg Ser Cys Leu Ala Arg Phe Gly Gly Gly Gly Ala  
 271 85 90 95

**RAW SEQUENCE LISTING**

PATENT APPLICATION: US/10/537,612

DATE: 03/01/2006

TIME: 13:55:57

Input Set : A:\Q88296 - 10-537,612 sequence listing.txt  
Output Set: N:\CRF4\03012006\J537612.raw

273 Asp Glu Gln Ala Asp Ile Cys Glu Ser Leu Ala Asp Trp Ala Asp Glu  
274 100 105 110  
276 His Tyr Arg Ser Cys Leu Ala Arg Phe Gly Asp Asp Gly Glu Asn Leu  
277 115 120 125

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/537,612

DATE: 03/01/2006

TIME: 13:55:58

Input Set : A:\Q88296 - 10-537,612 sequence listing.txt  
Output Set: N:\CRF4\03012006\J537612.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date